

Lecture 12 - High Availability (via Parallel Sysplex)

High Availability?

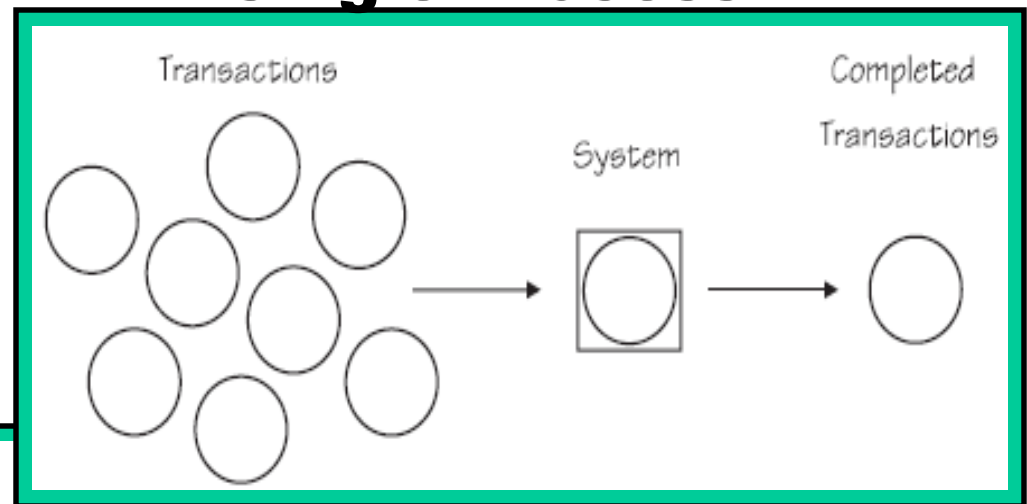
- If the system is “up” 99.9% of the time, it will be “down” 43 minutes/month (or over 8 hours/year)
- If the system is “up” 99.99% of the time, it will be “down” 4 minutes/month (or over 50 minutes/year).
- If the system is “up” 99.999% of the time, it will be “down” ~1 minute/month (or over ~5 minutes/year).
- How to achieve High Availability
 - » Need to account for hardware and software failures, so need to have redundant systems

Parallel Processing

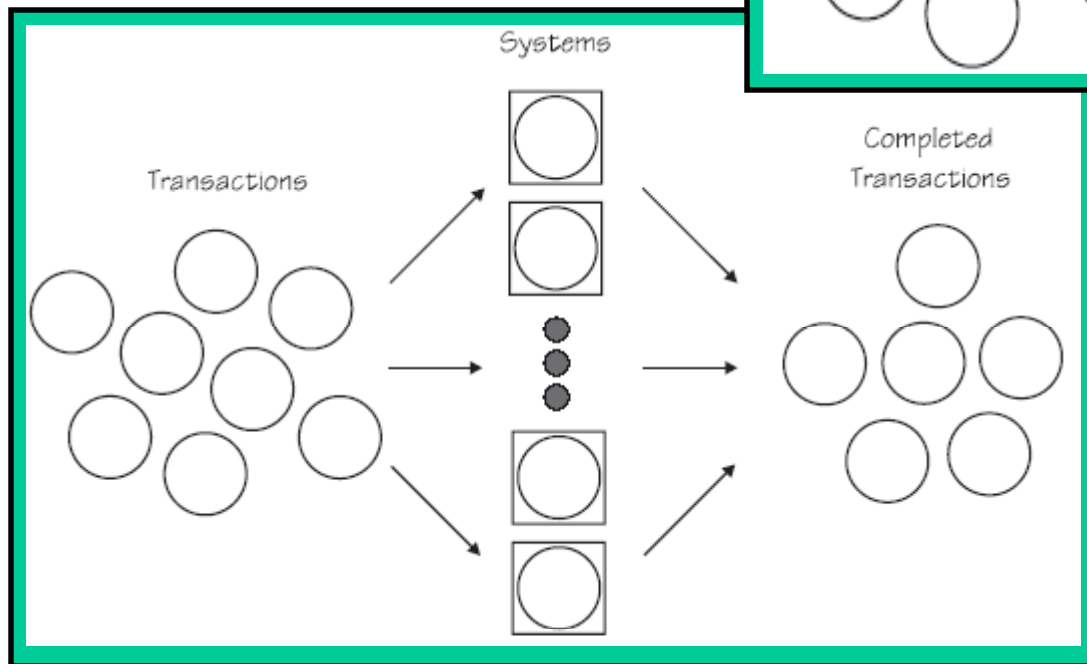
- Simultaneous processing of the same workload by multiple CPCs
- Processing a workload in parallel can improve response times:
 - longer running tasks do not “slow down” response time for shorter running tasks
 - Gain improve resource utilization within and across systems
- If used across systems, can also improve availability

Parallel Processing

Single Process



Parallel Processing

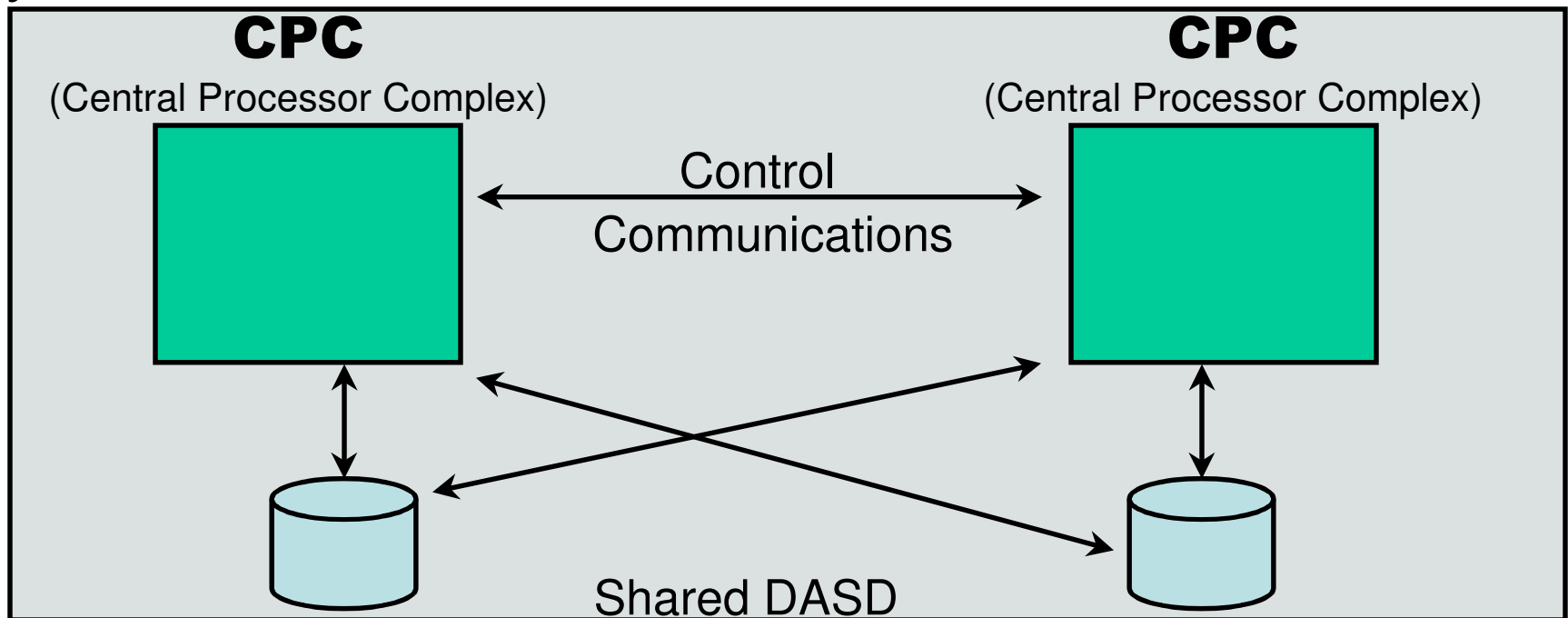


Redundant systems

- Achieved via clustering
- Can be thought of as a form of Parallel Processing:
 - Simultaneous processing of the same workload by multiple CPCs
- Types of clustering
 - Loosely coupled
 - Tightly coupled

Loosely-Coupled Complex

- Parallel Processing within one system (or across systems)
- Consists of several systems (images) sharing a work load queue.
- One of the images could process work using data which is shared with the other images
- Products and Applications need to use their own synchronization



Challenges with a Loosely-Coupled Complex

- Each application needs to communication (with the “same” application running in the other complexes)
- Synchronization between two images can be challenging, but if the number increases to 3 or 10:
 - The complexity is more challenging
 - The CPU time can increase dramatically
 - » Imagine trying to synchronize 10 database management systems, each needing to be fully and accurately coordinated across the 10 images (ex. the locking would need to be coordinated across the 10 images)

What is a Sysplex (SYStems ComPLEX)?

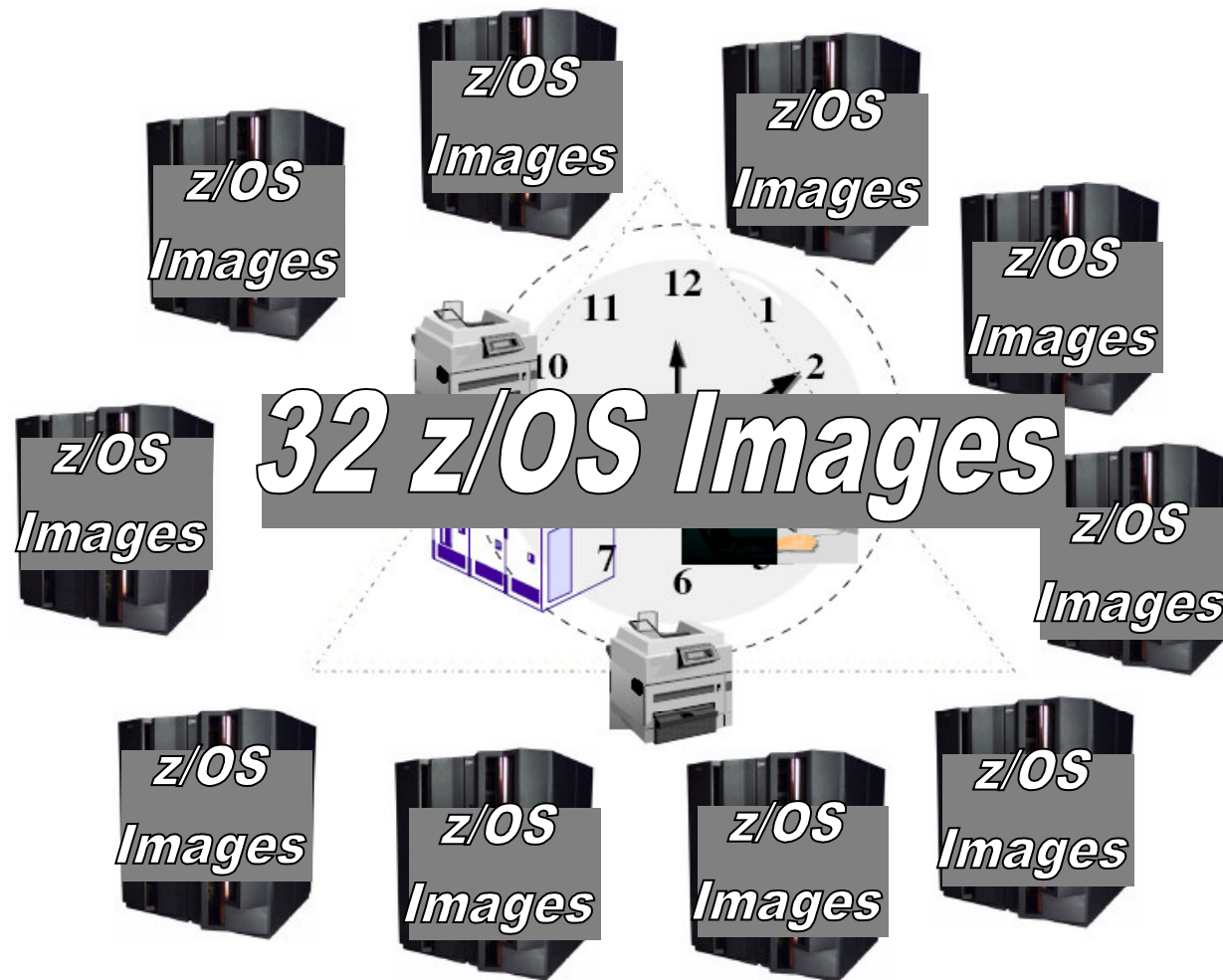
- A sysplex is a collection of z/OS systems (aka IBM mainframe CPCs) that cooperate, using certain hardware and software products, to process work

-- The goal of a Sysplex is to --

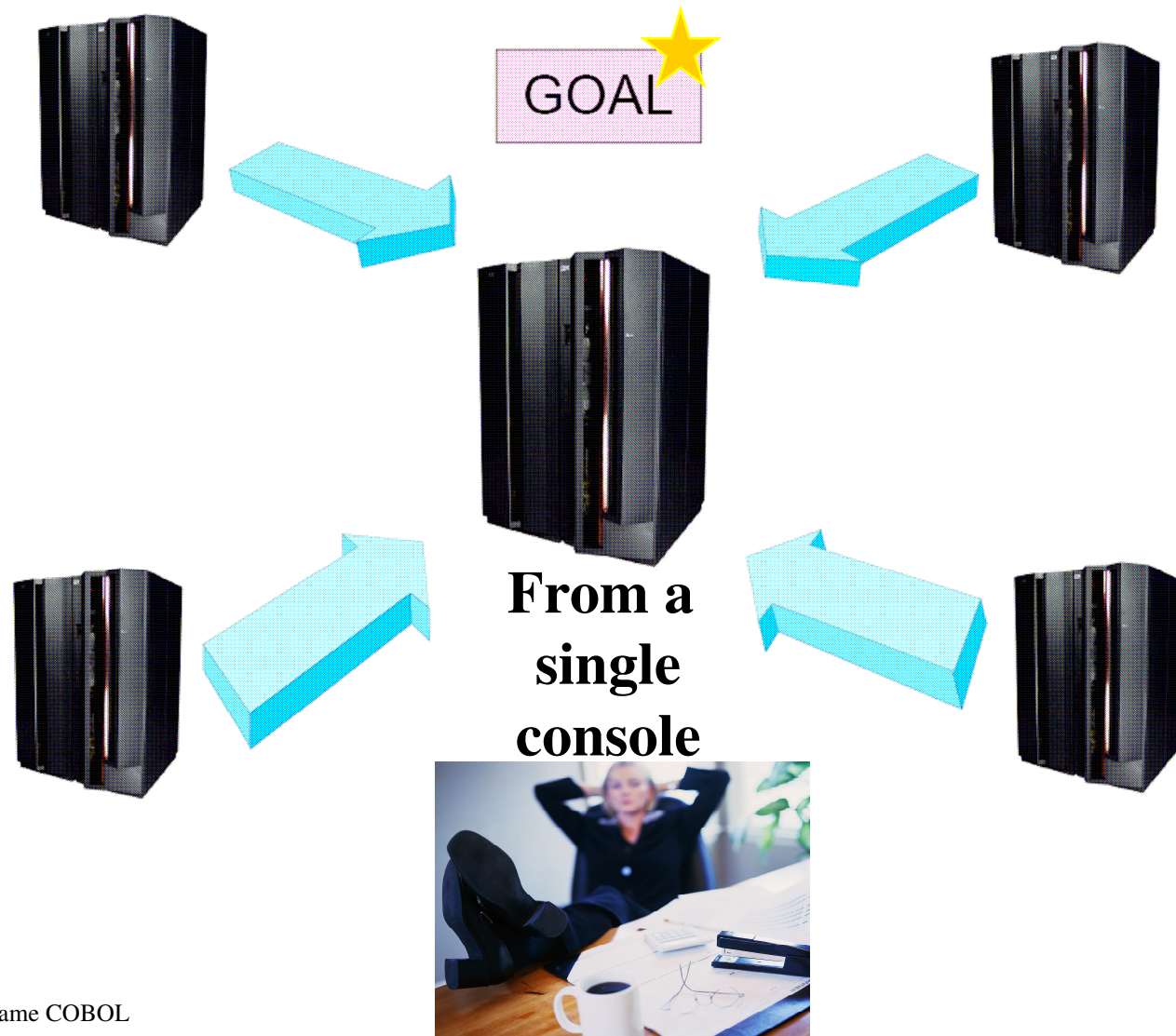
Make it easier to create loosely coupled parallel processing systems.

- **Benefits of a Sysplex include:**
 - Improved growth potential
 - Improved load balancing
 - Improved level of availability

SYStems comPLEX or SYSPLEX



Single System Image



When is using a Sysplex a good idea?

It will address any of the following types of work

- Large business problems that involve hundreds of end users, or deal with volumes of work that can be counted in millions of transactions per day.
- Work that consists of small work units, such as online transactions, or large work units that can be subdivided into smaller work units, such as queries.
- Concurrent applications on different systems that need to directly access and update a single database without jeopardizing data integrity and security.

What else a Sysplex can do for YOU !

- Platform for continuous availability so that applications can be available 24 hours a day, 7 days a week, 365 days a year
- **Ability to do more work**
 - Greater capacity
 - Improved ability to manage response time
 - Platform for further capacity and response time advances
- **Greater flexibility**
 - Ability to mix levels of hardware and software
 - Ability to dynamically add systems
 - An easy path for incremental growth
 - Varied platforms for applications, including parallel, open, and client/server
- **Workload balancing**

What is Parallel Sysplex?

- Technology that allows multiple mainframes (a *cluster*) to act as one.
- A Parallel Sysplex combines:
 - Data sharing (typically using Peer to Peer Remote Copy)
 - Parallel computing
- Parallel Sysplex allows:
 - A cluster of up to 32 computers to share a workload for high performance and high availability
 - Parallel Sysplex provides horizontal scaling.

What is parallel sysplex

- Builds on the strength of zSeries servers by linking up to 32 images to create the industry's most powerful commercial processing clustered system
- Every “cloned” application can run on every image
- Hardware and software can be maintained non-disruptively

What is a parallel sysplex?

- Every “cloned” application can run on every image
- Direct concurrent read/write access to shared data from all processing nodes
- No loss of data integrity
- No performance hit
- Transactions and queries can be distributed for parallel execution based on available capacity and not restricted to a single node

Continuous availability

- Within a parallel sysplex cluster, it is possible to construct an environment with no single point of failure
- Peer instances can of a failing subsystem can take over recovery responsibility for resources held by the failing instance
- Alternatively the failing subsystem can be automatically restarted on still healthy systems

Continuous availability

- In a parallel sysplex it is possible that the loss of a server may be transparent to the application and the server workload redistributed automatically with little performance degradation
- Each system is still individual
- Software upgrades can be rolled through one system at a time on a sensible timescale for the business

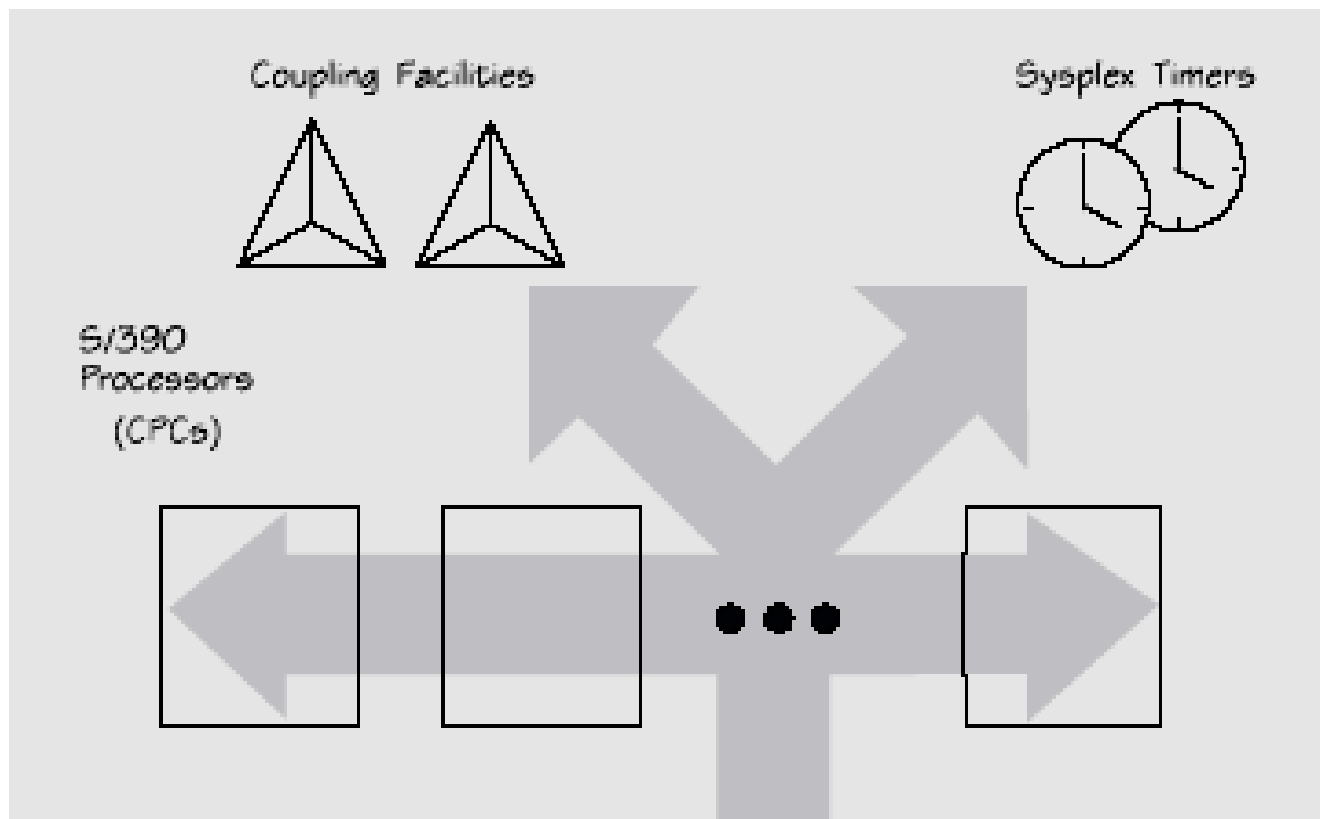
Single System Image

- The sysplex should appear as a single image to operator, end user, database administrator and others
- Single point of control
- Persistent single image across failures

Incremental Upgrades

- Replace / Upgrade each system (by taking a system down, upgrading, and then restarting that system).
- But what if the upgrade changed something such as a message format?
 - » Much easier to upgrade if system is upward compatible
 - » Some systems can support 10 or 20 year old code bases

Sysplex Component Overview



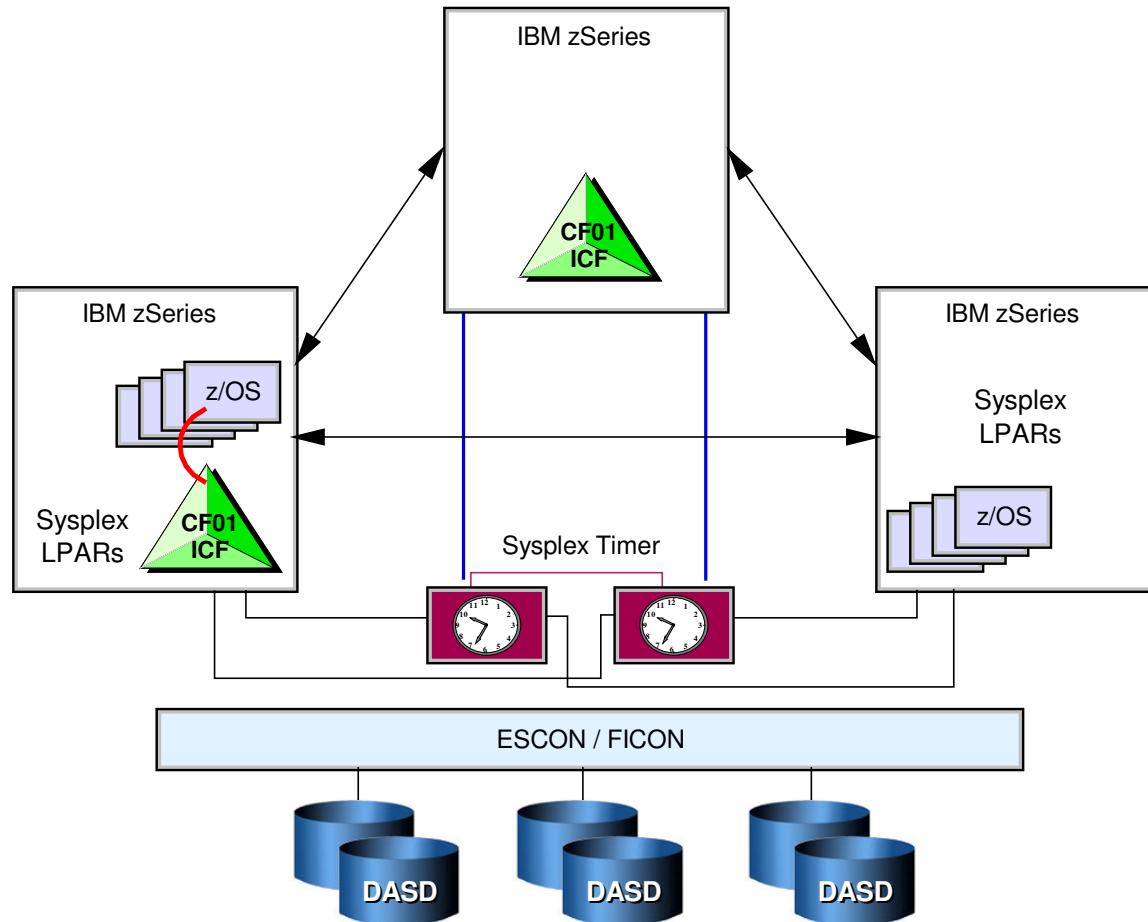
Components of a Parallel Sysplex environment include:

- **A Sysplex Timer:**
 - Synchronizes all member systems' clocks.
 - *Server Time Protocol (STP) replaced the Sysplex Timers in 2005*
 - *While the Sysplex Timers are physically separate machines, STP is an integral feature within the mainframe itself..*
- **Global Resource Serialization (GRS):**
 - Allows multiple systems to access the same resources concurrently
- **Cross System Coupling Facility (XCF):**
 - Allows high speed peer to peer communication (includes caching, locking, updating locked data, and workload balancing)
 - The Coupling Facility may be either a dedicated external system or ICFs (Internal Coupling Facilities within the mainframe).
- *With STP and ICFs it is common to construct a complete Parallel Sysplex installation solely with two mainframes and cabling*

Coupling Facility – Glue for Communication

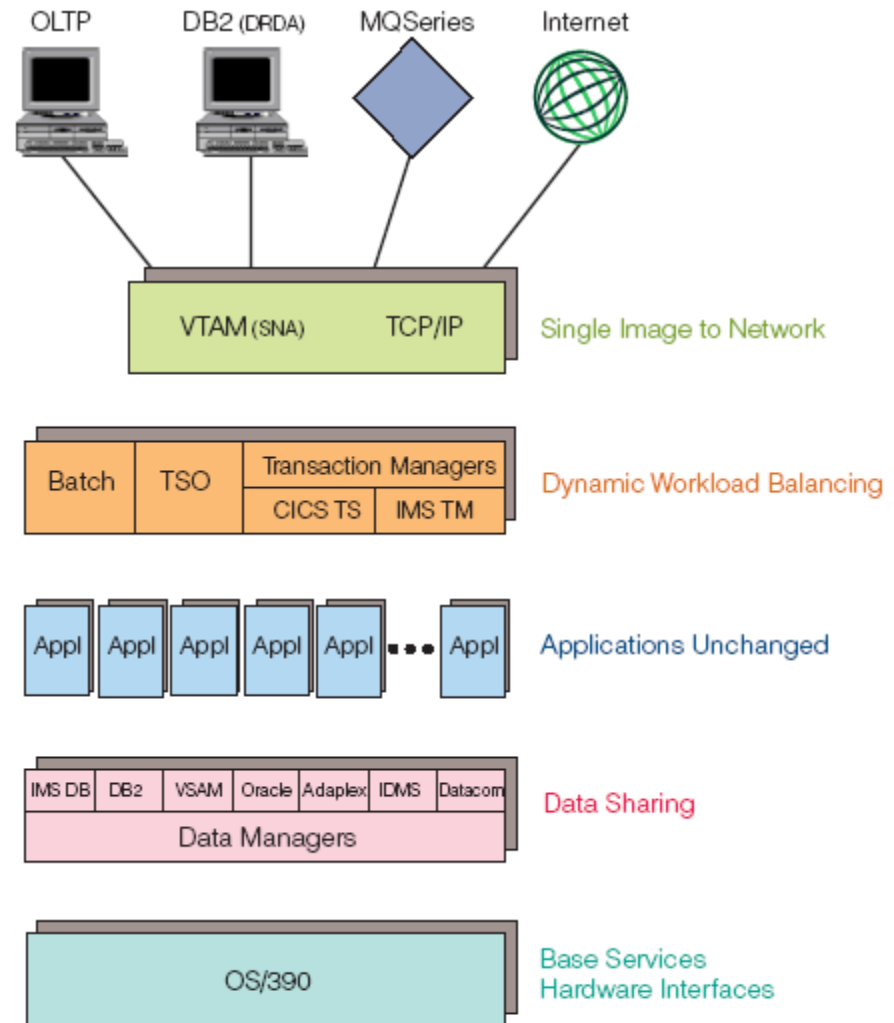
- Within the Coupling Facility, storage is dynamically partitioned into structures. z/OS services manipulate data within the structures. Each of the following structures has a unique function:
- **Cache structure**: Supplies a mechanism called *buffer invalidation* to ensure consistency of cached data. The cache structure can also be used as a high-speed buffer for storing shared data with common read/write access.
- **List structure**: Enables authorized applications to share data that is organized in a set of lists, for implementing functions such as shared work queues and shared status information.
- **Lock structure**: Supplies shared and exclusive locking capability for serialization of shared resources down to a very small unit of data.

Sysplex Parallel Hardware Overview



Parallel Sysplex Software Stack

IBM Software Adapted to run in a sysplex environment



Parallel Sysplex
High level view of the software stack

Symmetry - A Key to Parallel Sysplex

- **Symmetry is when:**
 - All the systems (images) “look/act the same”
 - Share the same resources
- In real life, all systems will not be exactly the same -- but Symmetry is required for a robust parallel sysplex environment
- **Symmetry will enable:**
 - Data Sharing
 - High Availability
 - Dynamic workload balancing
 - Incremental Upgrades

Data Sharing

Data sharing is required to enable:

- **Dynamic workload balancing** -- can more easily distribute work across systems
- **High availability** -- since all data can be accessed by all systems, the loss of one system will not impact the availability of the other systems

Dynamic workload balancing

- The entire parallel sysplex cluster can be regarded as a single logical resource to end users and business applications
- Work can be dynamically distributed to nodes with available capacity
- Workload balancing also allows running of diverse applications while maintaining response times critical to the business

Reasons for Asymmetry: Affinities

Reasons for Asymmetry: Affinities

- Certain I/O devices (ex. Print server) might only be available to a specific system
 - » Must be able to switch the I/O Device to another system
- Hardware might be different across the systems
- Software might be different across the systems (ex. Software license cost might require that certain software only run on certain hardware)
- Need to ensure the software license allows for the software to be run on another system as a “backup”

Reasons for Asymmetry: Affinities

- Badly developed software that does not allow multiple instances of the software to run at the same time.
 - An example would be global variables (ex. Filenames)
 - Since Parallel Sysplex is a fairly new concept (developed in the 1990's), much of the “older” software will not be able to utilize parallel sysplex technology...
- For new applications, **it should be a design goal to:**
 - Run in a Parallel Sysplex environment with no application changes

GDPS

- A geographically dispersed parallel sysplex is the ultimate disaster recovery and continuous availability solution for a multi-site enterprise
- Two sites up to 100 fiber kilometers apart may be connected for synchronous updates
- Asynchronous techniques may be used over this distance

Parallel Sysplex v.s. Shared Memory Processor (SMP)

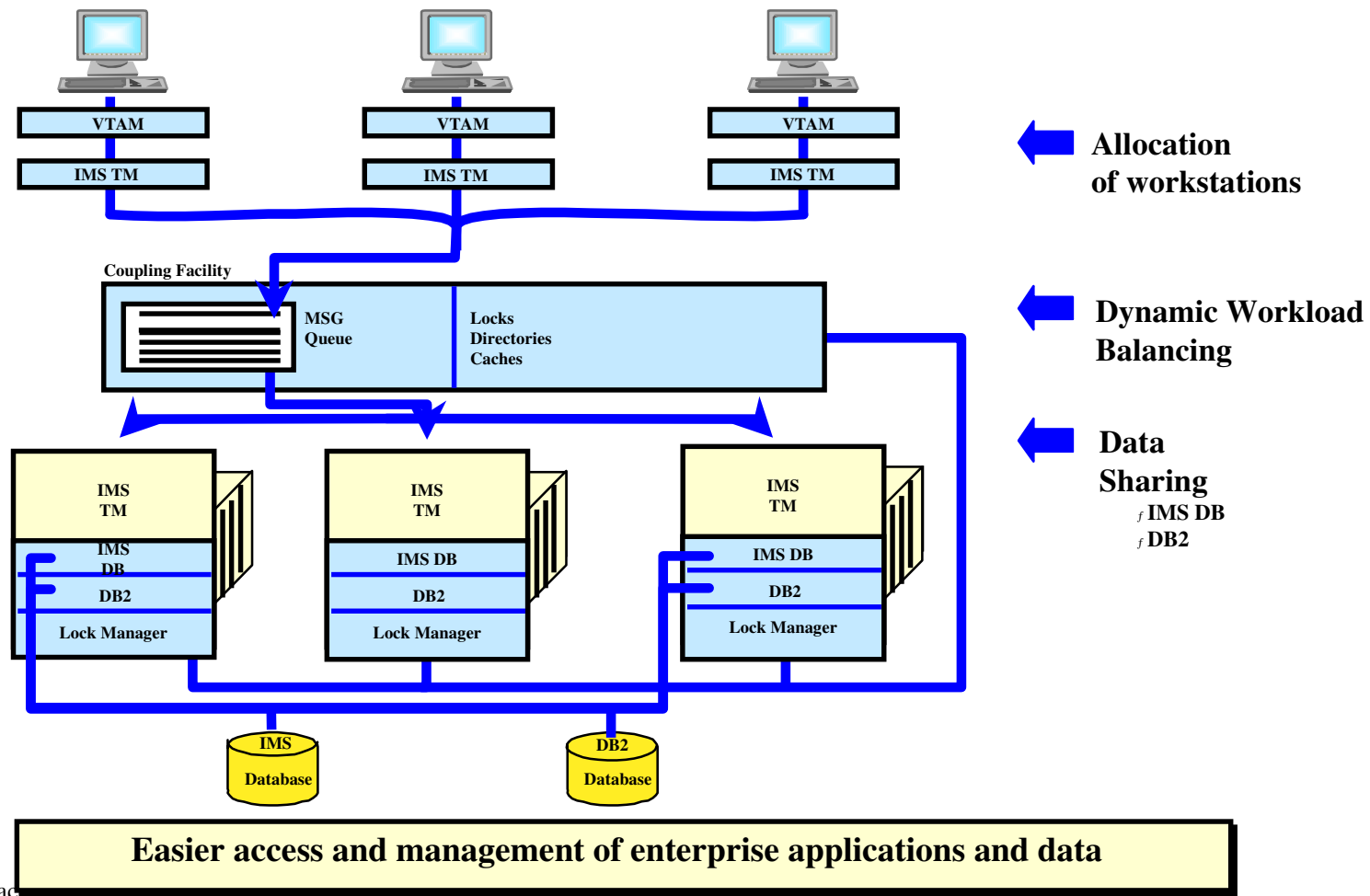
“*re*THINK” Application Design Choice

- ☐ Affinities
- ☐ In Memory Counter
- ☐ Processor local timestamp is not unique
- ☐ Single Record Database
- ☐ Parallelism of execution may have consequences
- ☐ Do not design to software levels

Applicability to a credit card transaction processing system

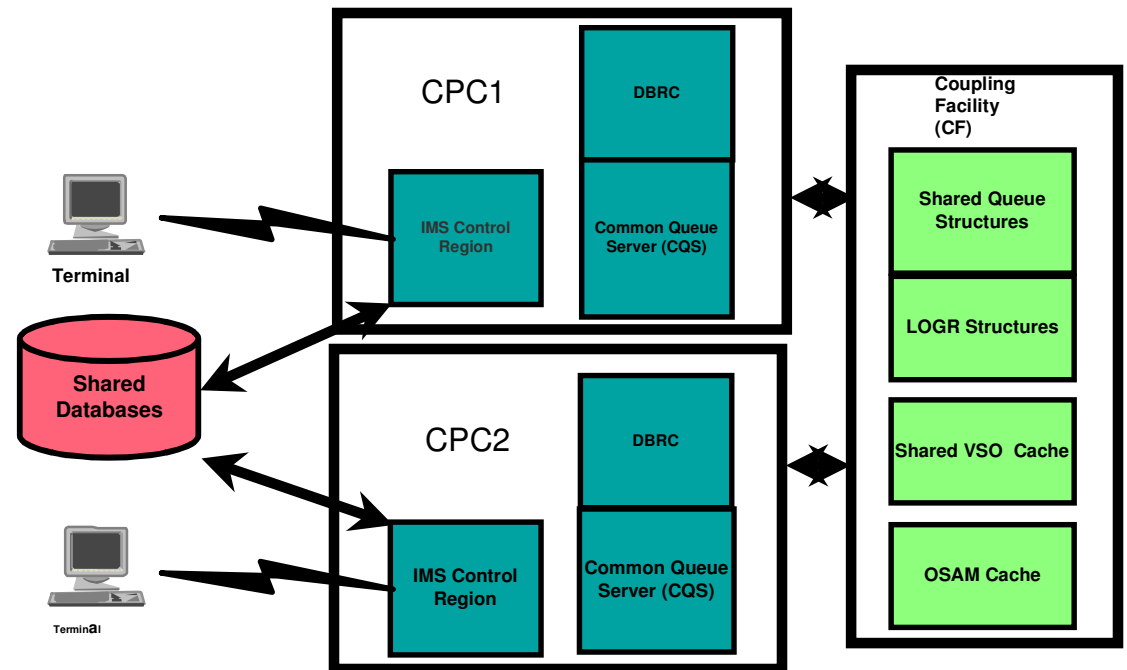
- What functions would benefit from running “in a parallel sysplex environment
- What might be some challenges to enabling the application to be run in the sysplex environment

IMS in a Parallel Sysplex Environment

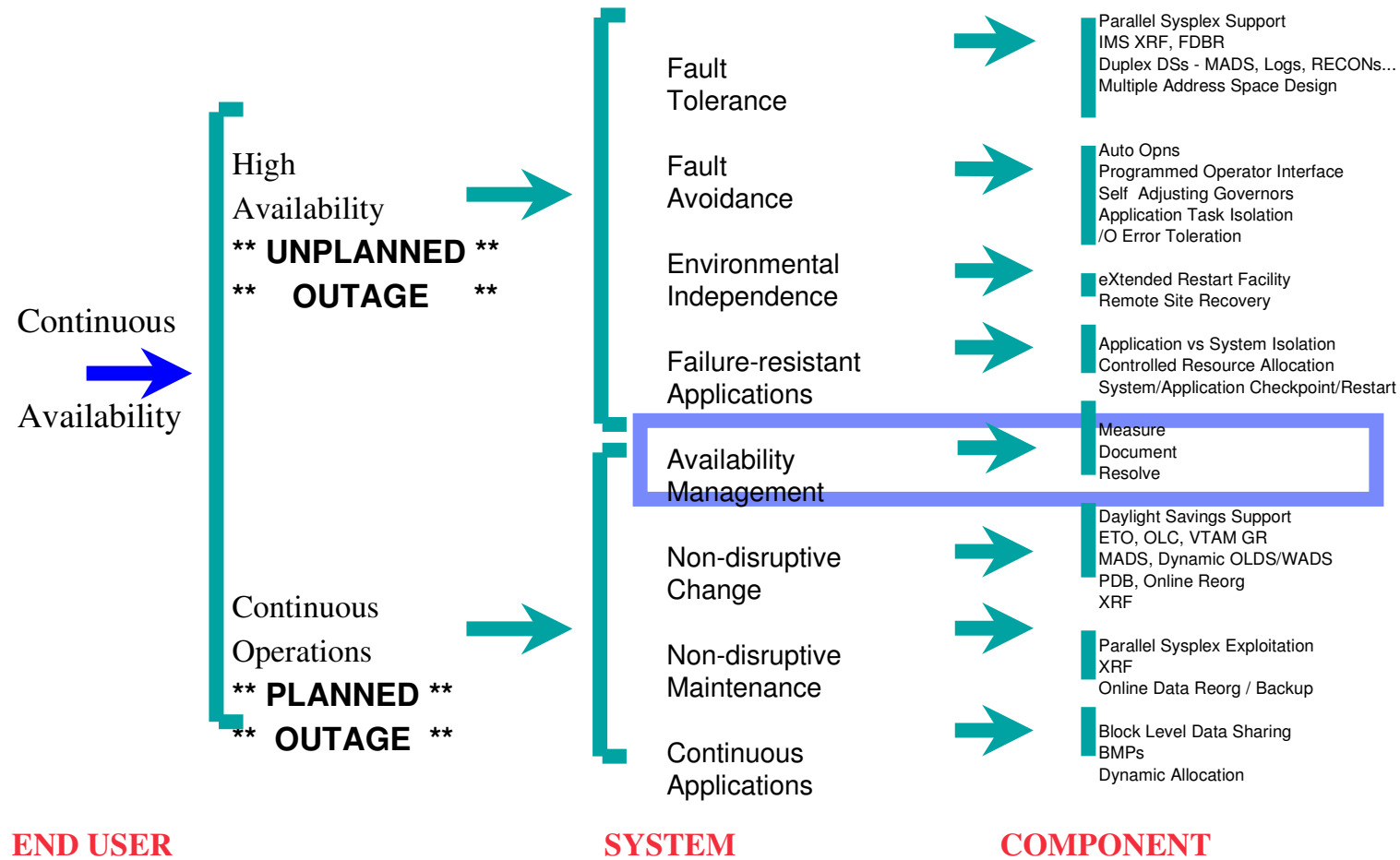


IMS in a Parallel Sysplex - Manageability

- PRESENT A SINGLE SYSTEM IMAGE
- PROVIDE EASE OF USE THROUGH A SINGLE POINT OF CONTROL ACROSS THE SYSPLEX
- ENABLE USERS TO RESUME STATUS ON ANOTHER IMS IN IMSPLEX
- COORDINATE/MANAGE ONLINE CHANGE ACROSS THE IMSPLEX



Ensuring IMS Availability



Parallel Sysplex Summary: Horizontal Scaling and High Availability

- Loosely coupled multiprocessing
- Hardware/software combination
- Requires:
 - Data sharing & Locking
 - Cross-system workload dispatching
 - Synchronization of time for logging, etc.
 - High-speed system coupling
- Hardware:
 - Coupling Facility
 - Sysplex Timer – Time Of Day clock synchronization
- Implemented in **z/OS*** and subsystems
 - Workload Manager in z/OS
 - Compatibility and exploitation in software subsystems, including **IMS*, VSAM*, RACF*, VTAM*, JES2***, etc.
- Also allows Rolling Maintenance System and Application Code